Protein Engineering vol.6 no.4 pp.397–408, 1993

Structure prediction of subtilisin BPN′ mutants using molecular dynamics methods

Andreas P.Heiner[1], Herman J.C.Berendsen[3] and Wilfred F.van Gunsteren[2]

Laboratory of Physical Chemistry and BIOSON Institute, University of Groningen, The Netherlands

[1]Present address: VTT/Bio, PO Box 202, SF-02151 Espoo, Finland

[2]Present address: Laboratorium für Physikalische Chemie, ETH-Zentrum, CH-8092 Zürich, Switzerland

[3]To whom correspondence should be addressed

In this paper we describe the achievements and pitfalls encountered in doing structure predictions of protein mutants using molecular dynamics simulation techniques in which properties of atoms are slowly changed as a function of time. Basically the method consists of a thermodynamic integration (slow growth) calculation used for free energy determination, but aimed at structure prediction; this allows for a fast determination of the mutant structure. We compared the calculated structure of the mutants Met222Ala, Met222Phe and Met222Gln of subtilisin BPN′ with the respective X-ray structures and found good agreement between predicted and X-ray structure. The conformation of the residue subject to the mutation is relatively easy to predict and is mainly determined by packing criteria. When the side chain has polar groups its exact orientation may pose problems; long-range Coulomb interactions may generate a polarization feedback involving system relaxation times beyond the simulation time. Changes induced in the environment are harder to predict using this method. In particular, rearrangement of the hydration structure was difficult to predict correctly, probably because of the long relaxation times. In all conversions made the changes observed in the environment were found to be history-dependent and in particular the hydrogen bonding patterns provided evidence for metastable substates. In all cases the structure predicted was compared with available kinetic data and the reduced activity could be explained in terms of changes in the configuration of the active site.

Key words: computer simulation/mutant structure/molecular dynamics/serine protease/solvent structure/subtilisin

Introduction

One of the main objectives of current protein chemistry is the design of proteins with specific functional properties. This 'protein engineering' (Craik *et al.*, 1985; Thomas *et al.*, 1985) requires, amongst others, computational prediction of both the structure of mutated proteins and their functional properties on the basis of these structural data. The reliability of the first step can be tested independently of the second step if crystallographic data for the mutants are available. Once both steps are mastered, protein engineering in the true sense will become possible.

In this article we address the question of how the first step (i.e. prediction of the structure of the mutant enzyme) can be made using molecular dynamics (MD) simulations. [For recent reviews on the methods and applications of MD we refer to van Gunsteren and Berendsen (1990a) or Karplus and Petsko (1990) and references therein.] The mutation can, in principle, be made abruptly. The conformation of the mutated residue may be modelled using a database of preferred side chain conformations (Janin *et al.*, 1978; Ponder and Richards, 1987), using sequence homologies (Doolittle, 1981), structural data derived from the database of known crystal structures (Bernstein *et al.*, 1977) or packing criteria (Ponder and Richards, 1987; Lee and Subbiah, 1991). To remove bad contacts and local strain, the mutation may be followed by some kind of regularization scheme such as energy minimization (EM) with a subsequent MD equilibration. Using such a protocol, the resulting configuration may well be trapped in a metastable substate, as the local solvent structure is not considered when the side chain conformation is modelled. A similar problem may be encountered when systematic search procedures are employed.

Another approach to obtaining the structure of a mutant enzyme originates from the free energy perturbation approach. In this procedure the system is described by a mixture of the potential functions of the wild type and mutant enzyme and by slowly changing the balance between the two potential functions the structure of the mutant enzyme is obtained from the structure of the wild type enzyme. This technique is known as the slow growth procedure (Berendsen *et al.*, 1985) and is illustrated in Figure 1 for a process in which the final state Y has two distinct minimum energy conformations. Which of the two final conformations is obtained depends on the behaviour of the system near the branching point. The slower the balance between the two potential functions is changed, the higher the probability that the branch leading to the low-energy conformation (B-branch) is found, as the barrier between the A-branch and the B-branch is still small. When the system is changed more rapidly, it might get stuck in the high-energy state. This assumes that some of the information of the final state is present at the branching point, hence the steepest gradient with respect to λ at this point is likely to be directed towards the path leading to the B-branch. The rate of change of the system should be so slow that each intermediate state can be considered as a small perturbation of the equilibrated previous structure. In general this requirement is tested by comparing the free energy differences of the forward change (system X → system Y) and backward change (system Y → system X). When the difference is small in all intermediate states, the conversion is assumed to be made reversibly. Unfortunately, the determination of a free energy difference is computationally much more expensive, since the underlying distribution function has to be determined. In order to determine the free energy difference within 10% for the process sketched in Figure 1, the complete area within contours of 2 kT has to be sampled. A more accurate determination requires that both branches have to be sampled (contours within, e.g. 4 kT). In the simple case of the hydration of methanol, sufficient sampling took at least 50 ps per intermediate state; the change itself could be made much faster (unpublished data). Thus, if one is able to convert the system from state X into state Y in a 'near

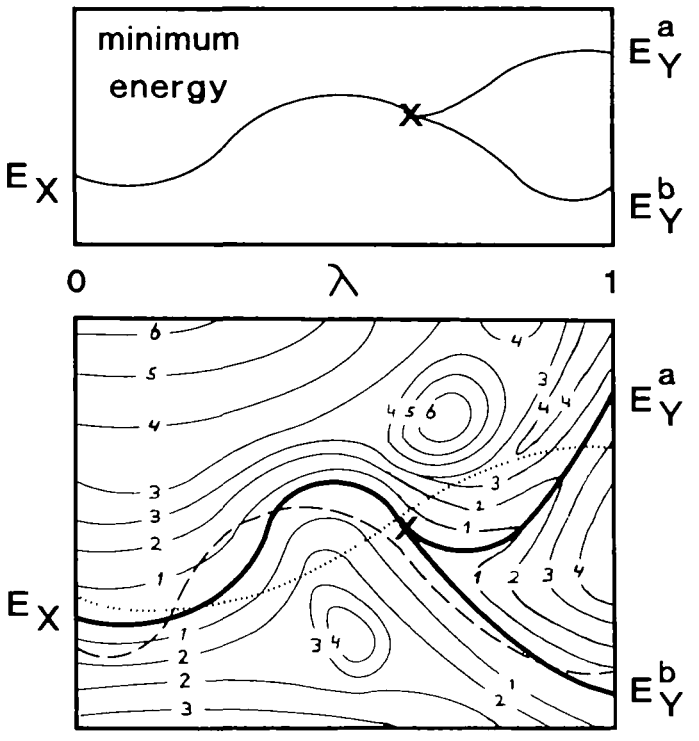

Fig. 1. Minimum energy curve (top) and energy contour plot (bottom) of the change of a hypothetical process as a function of the coupling parameter λ. The y-axis of the contour plot represents the $3N$ conjugate coordinates. Bold lines represent the local energy minimum of a cluster of configurations as a function of λ, thin lines are contours with energy 1 kT, 2 kT etc., relative to the absolute minimum energy at a given value of λ. The final state Y has two stable minima $E_Y^a > E_Y^b$ separated by an energy barrier >4 kT. The cross denotes the branching point; for this value of λ, the final conformation has been mixed into the system such that one distinct conformation (X-like system) splits into two distinct conformations (Y-like system). Near the branching point the energy barrier between the two distinct conformations is small. The broken and dotted lines represent two possible paths from the minimum energy conformation of the initial stage X into one of the local minimum energy conformation of the initial state Y.

equilibrium' manner without reference to the free energy difference, the change may be made much faster.

A validation of the predicted mutant structure can be obtained by comparing the average structure with available X-ray data. Furthermore, structural changes of the mutant enzyme with respect to the wild type enzyme may be used to explain altered kinetic activity. When the mutant enzyme is changed back into the wild type enzyme, the average structures before mutation and after backward mutation may be compared, provided that the total change forth-and-back is carried out in a time exceeding the intrinsic relaxation. Estimates of this correlation time range from 0.06 ps (Hermans *et al.*, 1992) to 3 ps (Berendsen, 1991). These correlation times are typical for arrangements within a local structure. There is much experimental evidence that the structural effects of point mutations in a protein are limited to such local structures (Bott *et al.*, 1985; Wells, 1990). Hydrogen bonding patterns during the initial equilibration period and the equilibration period after mutating forth-and-back may give clues whether the condition of a 'near equilibrium conversion' is met.

Another point that deserves some attention is the reliability of the force field and the computational procedure used. In a previous paper (Heiner *et al.*, 1992) we addressed this by analysing an MD run of native subtilisin BPN' in a crystal unit cell (referred to as the crystal simulation) and found good agreement with X-ray data. The largest structural deviations were

Table I. Crystallographic data and kinetic constants of the wild type enzyme subtilisin BPN' and three mutants at position Met222

Enzyme	Crystallographic data[a]		Kinetic date[b]		
	Resolution (Å)	*R*-factor	k_{cat} (s^{-1})	K_M (10^{-4} M)	k_{cat}/K_M (s/M)
Met	1.8	0.144	50	1 4	3.6×10^5
Ala	2.0	0.159	40	7.3	5.0×10^4
Gln	2.0	0.147	8	8.6	$9\,0 \times 10^3$
Phe	2.0	0.146	3 6	5.6	6.4×10^3

[a]Crystallographic data for Met from Bott *et al.* (1988) and Ala, Phe and Gln from Bott and Ultsch (1986).
[b]Activity with respect to *N*-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide (sAAPF-pna). Data for Met222 and Ala222 from Estell *et al.* (1985), for Gln222 from Bott *et al.* (1988) and for Phe222 from Bonneau *et al.* (1991).

found in the vicinity of Ca^{2+}-ions. Since the ion binding sites are rather distant from the active site, we do not expect that this will influence the results obtained here.

In this paper we will use the approach sketched above to predict the structure of a few mutants of subtilisin BPN' at the Met222 site. Upon oxidation of methionine the enzyme loses about 90% of its proteolytic activity (Stauffer and Etson, 1969). The enzyme can be made resistant to oxidation, but in general it loses activity in the process (Estell *et al.*, 1985). The Genentech/Genencor group has made all 19 possible variants of subtilisin BPN' at this site and the highly refined X-ray structures of a number of these mutants are known (Bott, 1985; Bott *et al.*, 1985, 1988). Kinetic and crystallographic data of mutants discussed are summarized in Table I. It should be noted that at the time of the calculations none of the mutant structures were available to us.

Method and computational procedure

As an initial molecular model we used the structure of native subtilisin BPN' at 1.8 Å resolution with an *R*-factor of 0.144 (Bott *et al.*, 1988). A periodic unit cell with four protein molecules and solvent molecules as described in an earlier paper (Heiner *et al.*, 1992) was used as a starting model. In order to reduce the computational effort, a sphere around the active site of one molecule was selected. All atoms belonging to charge groups whose centre of geometry was beyond 2.5 nm from this active site centre r_c.

$$r_c = \tfrac{1}{2}(r(\text{His64:N}_{\epsilon 2}) + r(\text{Ser221:O}_\gamma)) \qquad (1)$$

were removed from the system. Atoms belonging to charge groups of which the centre of geometry was in the outer 0.8 nm shell of the sphere were positionally restrained by a quadratic position restraining potential of which the force constant was given by

$$\begin{aligned} K_s(r_{ic}) &= 0 && r_{ic} < 1.7 \text{ nm} \\ K_s(r_{ic}) &= \frac{K_s(\infty)}{2}\left(1 + \sin\left(\pi \frac{r_{ic} - 1.9}{2.1 - 1.7}\right)\right) && \\ & && 1.7 \text{ nm} < r_{ic} < 2.1 \text{ nm} \\ K_s(r_{ic}) &= K_s(\infty) && 2.1 \text{ nm} < r_{ic} < 1.5 \text{ nm} \qquad (2) \\ & \text{with } K_s(\infty) = 9000 \text{ kJ/mol/nm}^2 & & \\ & \text{and } r_{ic} = \mid \mathbf{r}_i - \mathbf{r}_c \mid & & \end{aligned}$$

This way a smooth change from unrestrained to strongly positionally restrained atomic motion is achieved.

The GROMOS87 force field (van Gunsteren and Berendsen,

1987) for hydrated systems was used. Non-bonded interactions were calculated using the twin range concept (Berendsen *et al.*, 1986) with a short-range cut-off $R_{c,short}$ = 0.8 nm and a long-range cut-off $R_{c,long}$ = 1.2 nm. Both radii applied to the centres of geometry of charge groups to prevent the breaking of charge neutrality. The long-range Coulomb force was updated every 10 MD steps, as was the pair list used to calculate the short-range force. The spherical system was energy minimized for 1500 steps using the steepest descent method. Initial velocities were taken from a Maxwellian distribution at 298 K. The temperature was controlled by weak coupling of the system to a temperature bath (Berendsen *et al.*, 1984) (T = 298 K, τ = 0.1 ps, separate scaling for solute and solvent atoms). A leap-frog variant of the multiple time-step formalism (Teleman and Jönsson, 1986) was used for solute atoms only, with five intermediate steps per time step for the bond-stretching force. Time steps of Δt = 2.0 ps and Δt = 1.0 ps were employed. The system was equilibrated for 25 ps [During the first 20 ps of the equilibration, the bond lengths were constrained using the SHAKE-procedure (Ryckaert *et al.*, 1977; van Gunsteren and Berendsen, 1977)]. During the first 5 ps of the equilibration all atoms were positionally restrained to their initial positions by a quadratic restraining potential with a force constant K_i = 9000 kJ/mol/nm^2 that decreased linearly in time. For the shell atoms, the maximum of the two force constants K_i and $k_s(r)$ was taken.

Conversion of the native protein (system X) into a mutant protein (system Y) is done by making the atomic interaction function $V(r_1 \cdot r_N)$ a function of the so-called coupling parameter λ as well. The boundary conditions are given by $V(\lambda = 0) = V_X$, the potential energy function of the native protein and $V(\lambda = 1) = V_Y$, the potential energy function of the mutant enzyme. The path of conversion is determined by two quantities.

(i) The atomic map, i.e. the list of atoms whose properties are chosen to depend on the coupling parameter λ.
(ii) The parameterization of the potential $V(\lambda)$.

Atoms of the atomic map that have to be eliminated (e.g. the methyl group C_ϵ of Met222 in the conversion from native enzyme to the Ala-mutant) are made dummy atoms (i.e. atoms with no van der Waals or Coulomb interaction and a mass of 1 a.m.u.) in the final state. Atomic masses were scaled linearly in λ. Bonded interaction parameters involving one or more dummy atoms were kept at their values in the native system, but with zero force constant. When atoms had to be created (e.g. the C_ζ-atom in the conversion from the native protein into the Phe-mutant), dummy atoms were added to the methionine side chain such that a reasonable chemical structure (correct bond lengths, reasonable bond angles) of the mutant was generated, but without interactions with the environment. The conversion of native protein into the Gln-mutant allows for two atomic maps due to the asymmetric side chain of glutamine; both paths were used. A detailed description of the atomic map is given in Table II.

The parameterization of the hybrid potential $V(\lambda)$ was fixed during all conversions made and is given by (van Gunsteren, 1988, 1989).

$$V(r_1 \ldots r_N,\lambda) = \sum_{\text{bonds } i} V(b_i,\lambda) + \sum_{\text{angles } i} V(\theta_i,\lambda) + \sum_{\substack{\text{improper}\\ \text{dihedrals } i}} V(\xi_i,l) + \sum_{\text{dihedrals } i} V(\phi_i,\lambda) + \sum_{\text{pairs } ij} V(r_{ij},\lambda) \quad (3)$$

$$\begin{aligned} V(b_i,\lambda) &= \tfrac{1}{2}((1-\lambda)K_{b_i}^X + \lambda K_{b_i}^Y) \times (b_i - ((1-\lambda)b_{0,i}^X + \lambda b_{0,i}^Y))^2 \\ V(\theta_i,\lambda) &= \tfrac{1}{2}((1-\lambda)K_{\theta_i}^X + \lambda K_{\theta_i}^Y) \times (\theta_i - ((1-\lambda)\theta_{0,i}^X + \lambda \theta_{0,i}^Y))^2 \\ V(\xi_i,\lambda) &= \tfrac{1}{2}((1-\lambda)K_{\xi_i}^X + \lambda K_{\xi_i}^Y) \times (\xi_i - ((1-\lambda)\xi_{0,i}^X + \lambda \xi_{0,i}^Y))^2 \end{aligned} \quad (4)$$

$$V(\phi_i,\lambda) = (1-\lambda)K_{\phi_i}^X(1 + \cos(n_i^X \phi_i - \delta_i^X)) + \lambda K_{\phi_i}^Y(1 + \cos(n_i^Y \phi_i - \delta_i^Y))$$

$$V(r_{ij},\lambda) = (1-\lambda)(V_{vdW}^X(r_{ij}) + V_{Coul}^X(r_{ij})) + \lambda(V_{vdW}^Y(r_{ij}) + V_{Coul}^Y(r_{ij}))$$

where q_i is a force field parameter (bond length, bond angle etc.) $q_{i,0}$ the equilibrium value of q_i, and $K_{q_i}^X$ and $K_{q_i}^Y$ the force constant of interaction q_i in systems X and Y respectively. For dihedrals, $n_{q_i}^X$ and $n_{q_i}^Y$ are the multiplicities of dihederal i in systems X and Y and $\delta_{q_i}^X$ and $\delta_{q_i}^Y$ are the phase shifts of the dihedrals in the respective systems.

Once the atomic map and the λ-parameterization of the atomic interaction function are chosen, the time dependence of the coupling parameter λ determines the rate of change of system X into system Y. To quantify the rate of change as a function of λ is difficult, due to the complexity of the hybrid potential [equations (3) and (4)]. Thus, a linear time dependency is the most unbiased choice. However, it has been argued that a change in the van der Waals radius is the predominant factor in the system change. In order to keep the rate of change of the system constant in such a case, an nth power [n = 5 (van Gunsteren, 1990b), n = 6 (Postma, 1985), n = 11 (Mazor and Pettitt, 1991)] time dependence for λ has been suggested.

The quality of the predicted structure is assessed by comparing the average structure of the mutant protein with the mutant X-ray structure (process X → Y). Next the mutant protein is equilibrated for some time (usually 10 ps) before it is converted back into the native enzyme (process Y → X); average structures are determined. The latter allows for a test whether the changes were made in a reversible way (i.e. the substates along the path are 'near equilibrium'). The exact protocols are given in Tables II and III and Figure 2.

Comparison of the structures was done by calculating the root mean square positional difference (r.m.s.PD) of residues closest to the methionine side chain, i.e. residues His64, His67, Tyr217, Ser221–Ser224 and the backbone of residues Asn218–Thr220 (Figure 3a). These residues partly shield the methionine from the solvent. Residues Met222 and Tyr217 are part of the S_1' binding site [nomenclature of Schechter and Berger (1967) is used]. The r.m.s.PD was calculated using a translational/rotational fit of the C_α-atoms onto the reference structure; the fit applied to all residues whose C_α-atom was within 1.5 nm of the active site centre [equation (1)] at the start of the equilibration period.

The root mean square positional fluctuation (r.m.s.PF) was calculated with respect to the average structure; no fit procedures were applied.

The definition of a hydrogen bond is in accordance with Köhler *et al.* (1987); when the acceptor–hydrogen distance is less than 0.25 nm and the donor–hydrogen–acceptor angle is larger than 135°, a hydrogen bond is observed. The donor is the first atom mentioned.

Table II. Atomic map of the three conversions of Met222 into one of its mutants

Mutant	X-state	Y-state	Atom X-state	Atom Y-state
Ala	Met	Ala	C_γ	Dummy
			S_δ	Dummy
			C_ϵ	Dummy
Phe[a]	Phe	Met	C_γ	C_γ
			$C_{\delta 1}$	S_δ
			$C_{\epsilon 1}$	C_ϵ
			$C_{\epsilon 2}$	Dummy
			C_ζ	Dummy
Gln1[b]	Gln	Met	C_δ	S_δ
			$O_{\epsilon 1}$	C_ϵ
			$N_{\epsilon 2}$	Dummy
			$H_{\epsilon 21}$	Dummy
			$H_{\epsilon 22}$	Dummy
Gln2[b]	Gln	Met	C_δ	S_δ
			$O_{\epsilon 1}$	Dummy
			$N_{\epsilon 2}$	C_ϵ
			$H_{\epsilon 21}$	Dummy
			$H_{\epsilon 22}$	Dummy

[a]Dummy atoms are inserted such that the phenyl group is coplanar.
[b]Dummy atoms are inserted such that C_γ, S_δ, C_δ and dummy atoms are coplanar.
A dummy atom has zero van der Waals/Coulomb interaction parameters and a mass of 1 a.m.u. Bonded interactions involving one or more dummy atoms have zero force constant.
Atomic map of the three mutants. $V(\lambda = 0)$ is defined as the X-state, $V(\lambda = 1)$ is defined as the Y-state. Bonds involving dummy atoms are of correct length, angles involving dummy atoms are distorted.

Table III. Perturbation protocol used in the five conversion simulations

Mutant	Code[a]	Δt (fs)	t_{X-Y} (ps)	$T_{eq,Y}$ (ps)	t_{Y-X} (ps)	$t_{eq,X}$ (ps)	λ-time-dependency
Ala	A1	1.0	40.0	25.0	–	–	Linear[b]
	A2	2.0	10.0	10.0	10.0	10.0	Sixth root[c]
Phe	P1	2.0	12.0	10.0	12.0	10.0	Sixth power[d]
Gln1	G1	2.0	10.0	10.0	10 0	10.0	Sixth power[e]
Gln2	G2	2.0	10.0	–	–	–	Sixth power[e]

[a]Code as used in article.
[b]$\Delta\lambda$ linear over full conversion time (Figure 2, curve A1)
[c]$\Delta\lambda$ linear over intervals of 1 ps; interval size determined from sixth root of t/t_{X-Y}, where t is the time since the conversion started (Figure 2, curve A2).
[d]$\Delta\lambda$ determined from sixth power of t/t_{X-Y}, where t is the time since the conversion started (Figure 2, curve P1).
[e]$\Delta\lambda$ linear over intervals of 1 ps; interval size determined from sixth power of t/t_{X-Y}, where t is the time since the conversion started (Figure 2, curve G1).

Results

Wild type enzyme Met222

Since the structural differences between the native enzyme and the mutant enzymes are expected to be small (Bott *et al.*, 1986; Alber *et al.*, 1987), it is of major importance to first analyse structural differences between the average native structure and the X-ray structure, induced by force field or methodology. Hereto the wild type enzyme was simulated for 10 ps following the equilibration period.

The r.m.s.PD of the average structure with respect to the X-ray structure of the backbone atoms near the mutant residue (Met222) is small (0.032 nm), indicating a similar backbone folding (Figure 3b). The r.m.s.PD of the side chain atoms is considerably larger, 0.105 nm (Table IV). The major contribution (0.27 nm) comes from the Tyr217−phenol group, which shows at 44° rotation around the $C_\beta - C_\gamma$ bond. The phenol group is pushed away by a solvent molecule in the S_1' binding crevice and a hydrogen bond between Tyr217:O_η and a water molecule in a water channel (Bott *et al.*, 1988; McPhalen and James, 1988) is formed. The water molecule in the binding crevice is not found in the X-ray structure, but existence of a solvent molecule in this area is well documented (Sielecki *et al.*, 1979). The average χ_2-angle found in the simulation (100°) is in good agreement with values found in the database of preferred side chain conformations (Janin *et al.*, 1978; Ponder and Richards, 1987). The χ_2-angle found in the X-ray structure (147°) represents a high-energy conformation that may be selected due to thermal disorder of the phenol group (mean B-factor 0.13 nm^2). Another difference (r.m.s.PD 0.06 nm) between X-ray structure and average conformation concerns the χ_1-angle of the active site serine Ser221: −59° in the simulated structure, and −23° in the X-ray structure. The latter is a high-energy conformation that is scarcely found in the database of preferred side chain conformations. Finally, a minor difference for the Met222:χ_3 angle is observed: −119° in the X-ray structure compared to −156° in the average simulated structure. No preferred orientation for this angle could be found. Together with the rotation of the Tyr217−phenol group these changes in structure cause an enlargement of the volume accessible to the P_1' leaving group. While MD structures may deviate from the X-ray structure as a result of force field errors, the deviations of the dihedral angles near the active site may well be caused by the presence of the inhibitor phenylmethylsulphonylfluoride (PMSF) in the X-ray structure.

The hydrogen bonding patterns in the simulated structure are similar to those observed in the crystal simulation. The water molecule in the oxy anion hole and the hydrogen bonding network near the active site are well preserved.

Whether the system could be considered as equilibrated was tested by a consecutive simulation of 10 ps. The hydrogen

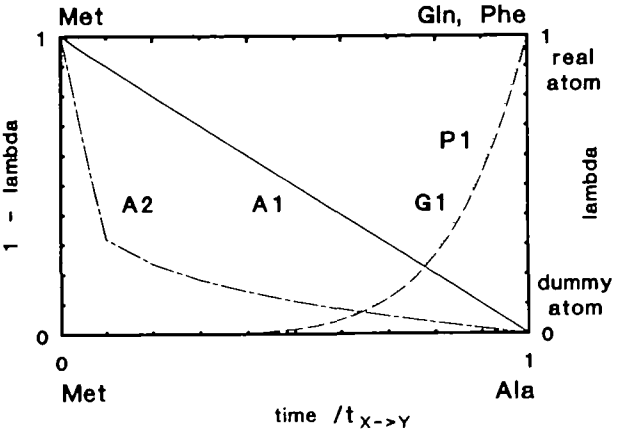

Fig. 2. Time dependence of the coupling parameter λ and the amount of real/dummy atoms as a function of time. The conversion time t_{X-Y} is in Table III. For the Met → Ala conversion (marked A1 and A2; left abscissa, 1 1 λ) real atoms are slowly changed into dummy atoms as a function of time. For the Met → Phe/Gln conversions (marked P1 and G1; right abscissa λ) dummy atoms are slowly changed into real atoms as a function of time.

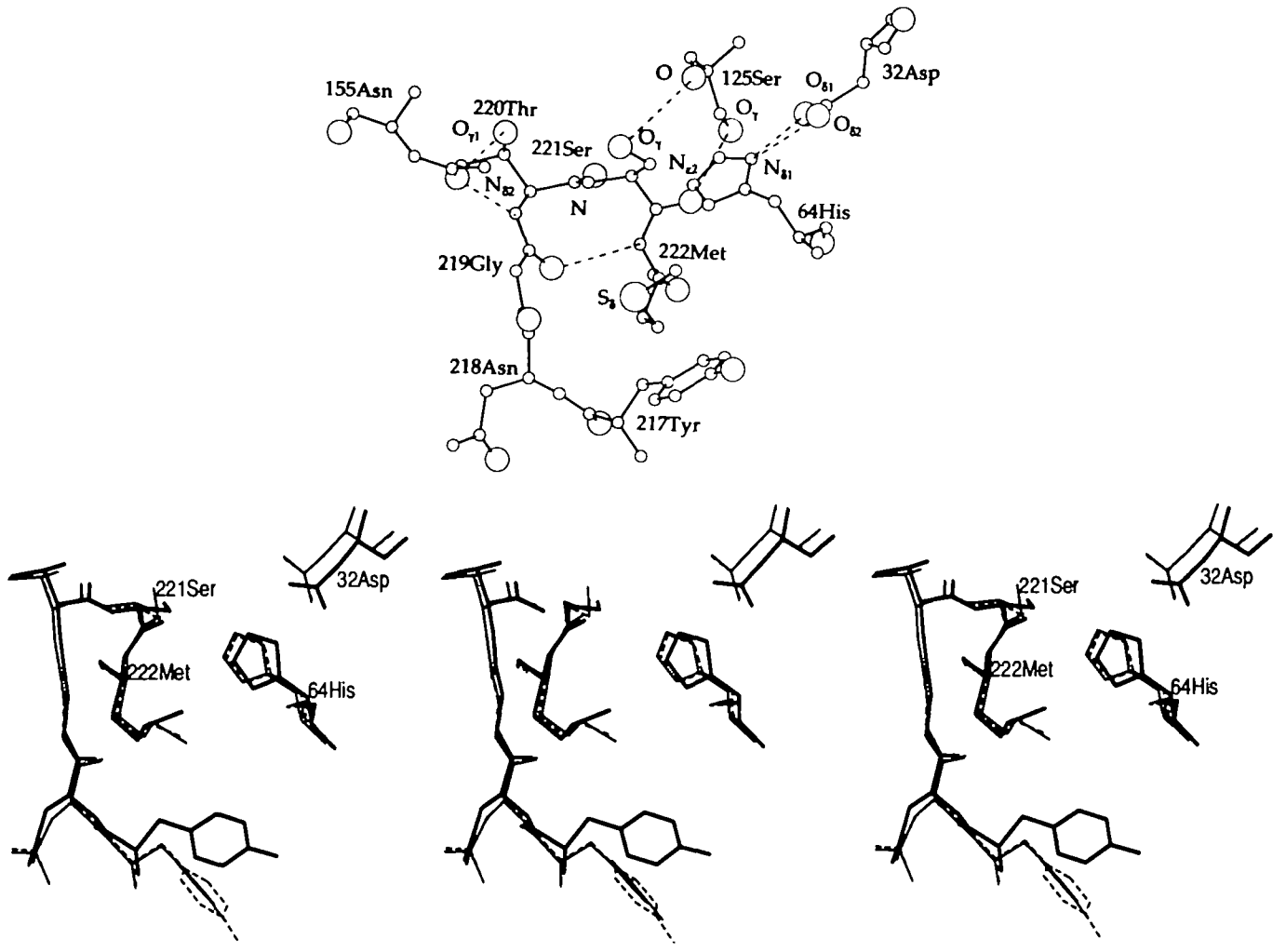

Fig. 3(a). Environment of residue Met222 in native subtilisin BPN' in the X-ray structure. Residues and atoms discussed throughout the paper are labelled; broken lines denote hydrogen bonds. (b) Environment of the residue Met222 in native subtilisin BPN' in the X-ray structure (bold lines) and the two 10ps average structures (20–30 ps: thin-dashed lines; 30–40 ps: thin lines). The active site residues His64, Ser221 and the residue subject to mutation Met222 are labelled.

Table IV. Quality of the predicted structure of Met222 mutants after conversion into Ala, Phe or Gln

	X-ray	Met	Ala(A2)	Phe(P1)	Gln(G1)
r m.s.PD (nm)					
Backbone	0.032[a]	0.011[b]	0.018	0.022	0.016
Side chain	0 105[a]	0.024[b]	0.034	0.097	0.047
r.m.s.PF (nm)					
Backbone		0.036	0.032	0.030	0.036
Side chain		0.046	0.045	0.035	0.046
Dihedral (°)					
Ser221:χ_1	− 23	− 60 (9)	5 (41)	− 61 (8)	62 (11)
Met222:χ_1	− 75	− 72 (9)	− 73 (10)	− 72 (13)	− 68 (6)
Met222:χ_2	73	70 (10)	71 613)	69 (10)	64 (10)
Met222:χ_3	−119	−156 (22)	−167 (22)	−164 (22)	−152 (31)
Distance (nm)					
His64:$N_{\epsilon 2}$−Ser221:O_γ	0.380	0.397	0.384	0.364	0.320
Asn155:$n_{\delta 2}$−Ser221:N	0.470	0.466	0.324	0.603	0.526
Ser221:O_γ	0.472	0.440	0.428	0.507	0.514
His64:$N_{\epsilon 2}$	0.729	0.713	0.628	0.673	0.741

[a]r.m.s.PD of the average structure from 20 to 30 ps with respect to the X-ray structure.
[b]r.m.s.PD of the average structure from 20 to 30 ps with respect to the average structure of the next 10 ps.
Indicators of the quality of the predicted structures of the Met()Y()Met mutants, i.e. after conversion of Met222 into Y222 into Met222. The r.m.s.PD is taken with respect to the average native structure from 20 to 30 ps. Averages over 10 ps.

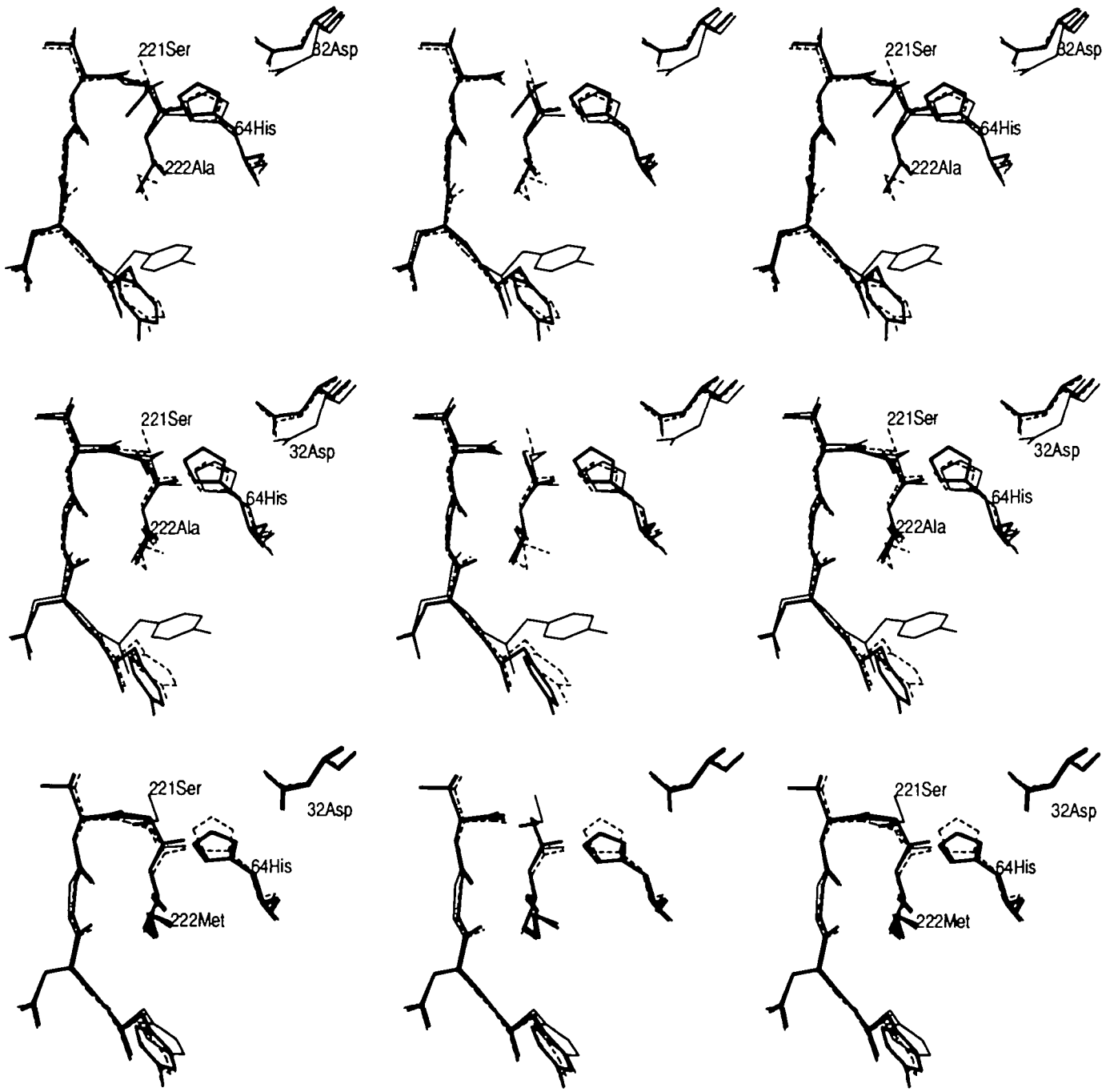

Fig. 4.(a). The average structure of the Ala-mutant using protocol A1 is shown (bold lines, <Met(A1)Ala>). The average structure of the native enzyme (thin-dashed lines, <Met>) and the X-ray structure of the Ala-mutant (thin lines) are shown as a reference The residues His64 and Ala222 as well as the hydroxyl group of Ser221 are labelled. Comparison of the solvent positions is not possible due to the diffusive behaviour of the molecules. (b) The average structure of the Ala-mutant (bold lines, <Met(A2)Ala>) using protocol A2. The average structure of the native enzyme (thin-dashed lines, <Met>) and the X-ray structure of the Ala-mutant (thin lines) are also shown. The residues His64 and Ala222 as well as the hydroxyl group of Ser221 are labelled. (c) The average Met-structure after mutating residue Met222 into Ala222 using protocol A2 (bold lines, <Met(A2)Ala(A2)Met>). In thin-dashed lines the average starting configuration <Met(A2)Ala> and in thin lines the average native structure <Met> are shown for comparison. The residues His64 and Met222 as well as the hydroxyl group of Ser221 are labelled.

bonding patterns and occupancies of the two simulations were almost identical; the r.m.s.PF of the two simulations of 10 ps were of similar magnitude (0.03 nm for 20–30 ps, 0.036 nm for 30–40 ps for backbone atoms; 0.046 nm for 20–30 ps, 0.046 nm for 30–40 ps for side chain atom). The r.m.s.PD of the average structures with respect to one another was low (0.011 nm for backbone atoms, 0.024 nm for side chain atoms). This indicates that the system may be considered equilibrated; the respective values for the r.m.s.PD express the statistical variation between two consecutive 10 ps-averaged structures.

Mutant enzyme Met222 → Ala222

The average structure of the Ala222-variant obtained using the first perturbation scheme <Met(A1)Ala> is almost identical to the average structure of the native enzyme (Figure 4a). [Average structures are denoted by '<' and '>'. The various mutants are

coded as follows: initial state—protocol of conversion—final state. Thus, <X(P1)Y(P2)X> denotes the average structure with residue Z at position 222. This structure is obtained by converting residue X222 into residue Y222 using protocol p1. After equilibration of the structure having residue Y222, the conversion from Y222 into Z222 is made using protocol P2. When P2 and P1 are identical, the second part of the conversion protocol P1 is to be taken (see also Table III).] The r.m.s.PD of the backbone atoms is 0.028 nm and that of the side chain atoms is 0.055 nm, both with respect to the average native structure. The backbone deviations are evenly distributed over the molecule and are only marginally larger than expected from statistical fluctuations. The side-chain of Ser221 yields the largest contribution, 0.165 nm. This is caused by a 112° rotation of the χ_1-angle with respect to the average native conformation. The gauche$^+$/gauche$^-$ difference [the definition of gauche$^-$/gauche$^+$ (+60°/−60°, also g^-/g^+) is in agreement with Janin *et al.* (1978)] between the wild type enzyme and the mutant enzyme is also observed in the X-ray structure, but the Ser221:χ_1 is only 29° in the mutant (Table V).

The mechanism of the transition is sketched in Figure 5a. The nucleophilic hydroxyl group forms a hydrogen bond to a water molecule (S1) that is also hydrogen bonded to His64:$N_{\epsilon 2}$. The water molecule S2 is partly located in the oxy anion hole and also forms a hydrogen bond to the imidazole. Due to the thermal motion the water molecules move upwards and the hydrogen bond from Ser221:O_γ to S1 is replaced by a hydrogen bond to S2 (Figure 5b). When solvent molecule S2 moves back to the original site, the χ_1-angle flips over (Figure 5C). In the native enzyme this process is also observed, but the gauche$^-$ conformation is less stable due to the proximity of the methionine side chain. During the conversion from Met into Ala the bonded force constants reduce, thereby increasing the flexibility of the methionine side chain. Also the van der Waals radii of the methionine side chain atoms decrease (0.37 of the original values at the moment the transition occurs). This may render the gauche$^-$ conformation as stable as the gauche$^+$ conformation. Due to the reorientation of the side chain the hydrogen bonding pattern is slightly different. In the wild type enzyme Ser221:O_γ serves as a donor to Ser125:O; in the Ala-mutant this hydrogen bond is completely lacking.

Another difference between the crystal structures is a very mobile water (B-factor 0.38 nm^2) in the Ala-mutant at the position where the methyl group is located in the wild type structure. When the conversion has been completed the site is still unoccupied; it is only after 15 ps equilibration that a solvent molecule diffuses in from the solvent channel near Tyr216:O_η.

We tested whether the protein structure changed due to the reorganization of the solvent. The protein hydrogen bonding pattern occupancies of the first 10 ps (empty Met222:C_ϵ-site) and the next 10 ps (occupied Met222:C_ϵ-site) were very much the same. The same holds for the respective r.m.s.PF values (backbone atoms, 0.036/0.035; side chain atoms, 0.045/0.045) and the r.m.s.PD of the two average structures with respect to the average native structure (backbone atoms, 0.030/0.023; side chain atoms, 0.070/0.066).

To reduce the computational effort we performed the conversion from Met to Ala in 5 ps. Using a linear time dependency for λ, the dihedral transition was not observed. Using a sixth-root time dependency ($\lambda \propto t^{1/6}$), the dihedral transition was observed, though at a larger λ-value (0.95, compared to 0.15 in protocol A1). When the mutation was reversed, the gauche$^+$ conformation was not restored; additional equilibration did not accomplish this either. In both cases the conformation of the methionine side chain was restored very well.

Increasing the time step to $\Delta t = 2$ fs while keeping the number of steps the same (protocol A2) should allow for a better equilibration of the system. However, the expected dihedral angle rotation did not occur, either during the conversion or during the equilibration (Figure 4b). Also, no solvent molecule diffused into the empty Met222:C_ϵ-site. The r.m.s.PD of the average structure <Met(A2)Ala> (backbone atoms 0.031 nm, side chain 0.054 nm) was similar to the r.m.s.PD of <Met(A1)Ala>, both with respect to the average native structure.

After converting the alanine side chain back into the methionine side chain, the conformation of the latter was restored almost perfectly (Table IV, Figure 4c). The r.m.s.PD of the average structure <Met(A2)Ala(A2)met> with respect to the average native conformation was small (backbone 0.018 nm, side chain

Table V. Quality of the predicted structure of the Met222 Ala mutants

Angle (°)			
Ser221:χ_1	29	52 (10)	−54 (14)
Distance (nm)			
His64:$N_{\epsilon 2}$−Ser221:O_γ	0.380	0.409	0.311
Asn155:$N_{\delta 2}$−Ser221:N	0.487	0.454	0.430
Ser221:O_γ	0.474	0.492	0.356
His64:$N_{\epsilon 2}$	0.716	0.636	0.600

Indicators of the quality of the predicted structure for the Met22Ala mutant using two different pathways. Averages over 10 ps.

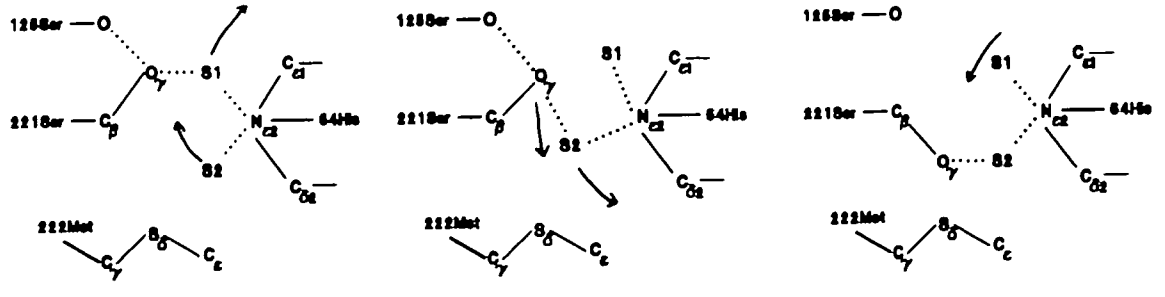

Fig. 5.(a). The predominant side chain conformation of the active site serine Ser221 is gauche$^+$. In this conformation the nucleophilic oxygen is hydrogen bonded to Ser125:O and a water molecule S1. The latter is also hydrogen bonded to His64:$N_{\epsilon 2}$. S2 is, mutually exclusive, hydrogen bonded to the oxy anion hole and His64:$N_{\epsilon 2}$. (Thin solid line, covalent bond; dotted line, hydrogen bond.) (b) Due to thermal motion the solvent molecule S2 replaces S1 in forming the hydrogen bond from Ser211:O_γ to His64:$N_{\epsilon 2}$. (c) Solvent molecule S2 moves back to the original position. Due to the hydrogen bond Ser211.O_γ−S2 the dihedral Ser211:χ_1 flips over to gauche$^-$. In the wild type enzyme the proximity of the methionine side chain makes this conformation less stable compared to the gauche$^+$ conformation. In the absence of the bulky side chain, the gauche$^-$ conformation may be as stable as the gauche$^+$ conformation sketched in Figure 5a.

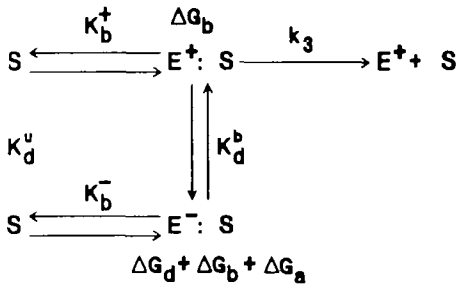

Fig. 6. Rate equations of the enzyme assuming a two-site model for the enzyme. The active form of the enzyme is the gauche$^+$ conformation for Ser221:χ_1. The difference in catalytic activity is caused by a difference in population of the gauche$^+$ and gauche$^-$ for the Ser221:χ_1 of the two enzymes ($\propto \exp(-\Delta G_d/RT)$). An additional energy difference ΔG_a is introduced to allow different gauche$^+$/gauche$^-$ ratios in the free and bound enzyme. Using this model, the binding affinity ΔG_b for sAAPF-pna for the wild type enzyme and the Ala-mutant is almost the same.

0.034 nm); only the Ser221:χ_1-angle is for 50% of the equilibration period in the gauche$^+$ conformation, the other 50% bring in the gauche$^-$ conformation.

The hydrogen bonding pattern in both the <Met(A2)Ala> and the <Met(A2)Ala(A2)Met> configurations are similar to the patterns found in the wild type enzyme; only the pattern near the active site is distorted. In particular, the nucleophilic oxygen is hydrogen bonded to His64:$N_{\epsilon 2}$ (69% occurrence) rather than to Ser125:O. The oxy anion hole is partly distorted due to a hydrogen bond from Asn155:$N_{\delta 2}$ to Ser221:O_γ. The hydrogen bond from His67:$N_{\epsilon 2}$ to Ile205:O is observed in the configurations having an alanine side chain at position 222 only.

The gauche$^-$ conformation of the active site serine is rather peculiar, since in almost all serine proteases the χ_1-angle of this residue is gauche$^+$. James *et al.* (1980) propose a two-site model for the nucleophilic oxygen in the unbound enzyme. Using this hypothesis, the X-ray data yield a gauche$^+$/gauche$^-$ ratio of 3:1 for the wild type enzyme and a ratio of 1:3 for the Ala-mutant. Large B-factors for the hdyroxyl group in both crystal structures support the hypothesis. For *Streptomyces griseus* protease A (SGPA) a gauche$^+$/gauche$^-$ ratio of 4 was found (James *et al.*, 1980). Data on porcine pancreatic elastase (Meyer *et al.*, 1988) show large B-factors for the hydroxyl group, but not for the C_β-atom; this also indicates two possible orientations for the active serine side chain. In the crystal simulation (Heiner *et al.*, 1992) we found a gauche$^+$/gauche$^-$ ratio of 6.

To test this two-site model against (the few) kinetic data available, we propose that only the gauche$^+$ conformation is active; the gauche$^-$ conformation is not capable of forming a tetrahedral intermediate. This assumption is reasonable, since in almost all complexed enzymes a gauche$^+$ conformation is observed. In the bound enzyme the water molecule S2 is replaced by the substrate and the mechanism for the g$^+$/g$^-$ dihedral transition is blocked. Hence, it is reasonable to assume an additional energy difference ΔG_a in the gauche$^-$ complex (Figure 6). The catalytic rate k_3 now becomes 51.6 s^{-1}, and the model also implies a higher fraction gauche$^+$ in the complex than in the free enzyme (wild type, fraction g$^+$: p_{unbound} ~ 0.75, p_{bound} ~ 0.97; Ala-mutant, fraction g$^+$:p_{unbound} ~ 0.25, p_{bound} ~ 0.77). Assuming steady state, the K_b for the wild type enzyme is 1.1 × 10^{-4} M and for the Ala-mutant 1.8 × 10^{-4} M; the difference in binding energy of the two enzymes is only 0.5 kT (~ 1.2 kJ/mol). This is consistent with the experimental finding of Estell *et al.* (1986) that when the difference in volume of the P_1-residue and the S_1-site is large, the binding between substrate and enzyme is strong (Estell *et al.*, 1986; Graycar, 1991). The mobile solvent at the Met222:C_ϵ-site may be the reason that in the Ala-mutant the g$^-$-conformation is favoured over the g$^+$-conformation.

Table VI. Quality of the predicted structure of the Met222Phe mutants

	X-ray	<Met(P1)Phe>
Dihedral (°)		
Ser221:χ_1	−73	−61 (8)
Phe222:χ_1	−39	−54 (7)
Phe222:χ_2	−69	−56 (11)
Distance (nm)		
His64:$N_{\epsilon 2}$−Ser221:O_γ	0.329	0.369
Asn155 $N_{\delta 2}$−Ser221:N	0.481	0.592
Ser221:O_γ	0.504	0.486
His64:$N_{\epsilon 2}$	0.727	0.686

Indicators of the quality of the predicted structure of the Met222Phe mutants. Averages over 10 ps.

Using protocol A2, the dihedral rotation was not observed. The water molecule S2 (Figure 5b) has moved from the oxy anion hole into the (partially) empty Met222:C_ϵ-site before the hydrogen bond to Ser221:O_γ is made. Using the A1 or A2 1fs protocol the solvent molecule S2 is closer to the imidazole than to the oxy anion hole. Apparently very minor structural details are important when the conversion is made fast.

The rotation of Ser221:χ_1 during the equilibration of Met(A2)Ala(A2)Met is not according to the scheme sketched earlier (Figure 5a–c). Another mechanism could not be identified. During the equilibrium simulation of the final Met structure a hydrogen bond is observed from Ser221:O_γ to His64:$N_{\epsilon 2}$. Although this bond occurs in the complexed enzyme, its presence in the free enzyme is unlikely. The absence of a hydrogen bond from Ser221:O_γ to Ser125:O indicates a strained structure as well. In the wild type enzyme, the carbonyl of a nearby serine serves as a 'parking place' for the nucleophilic oxygen (Meyer *et al.*, 1988; Geller *et al.*, 1990, 1991).

Mutant enzyme Met222 → Phe222

The r.m.s.PD of the average structure of the Phe222-mutant <Met(P1)Phe> with respect to the average native structure is moderate (0.023 nm for backbone atoms, 0.039 nm for side chain atoms) considering the large group that is inserted (Figure 7a). The largest r.m.s.PD for the backbone atoms is found for the mutated residue (0.034 nm); they are pushed away from the centre of the S_1'-pocket. The φ-ψ angles of the peptide plane Ser221−Phe222 differ by 15° with respect to the average native structure. The axes of the 3_{10}-helix and the central α-helix are less aligned. The largest r.m.s.PD for the side chain atoms is observed for residue Tyr217 (0.089 nm); the χ_2-angle rotates over −38° relative to the average native structure. The net effect of these structural changes is an enlargement of the S_1' pocket to accommodate the phenyl group. The conformation of the mutated residue Phe222 compares very well with the X-ray structure of this mutant (Table VI). The dihedral angles differ by 14° only, which is of the order of the fluctuations of these angles.

The methionine side chain conformation in the <Met(P1)Phe(P1)Met> structure is close to the conformation in the average native conformation; the average dihedral difference is only 2° (Table IV). The average structure <Met(P1)Phe(P1)Met> compares very well with the average native structure (Figure 7b; r.m.s.PD of backbone atoms 0.021 nm, r.m.s.PD of side chain atoms 0.042 nm). The

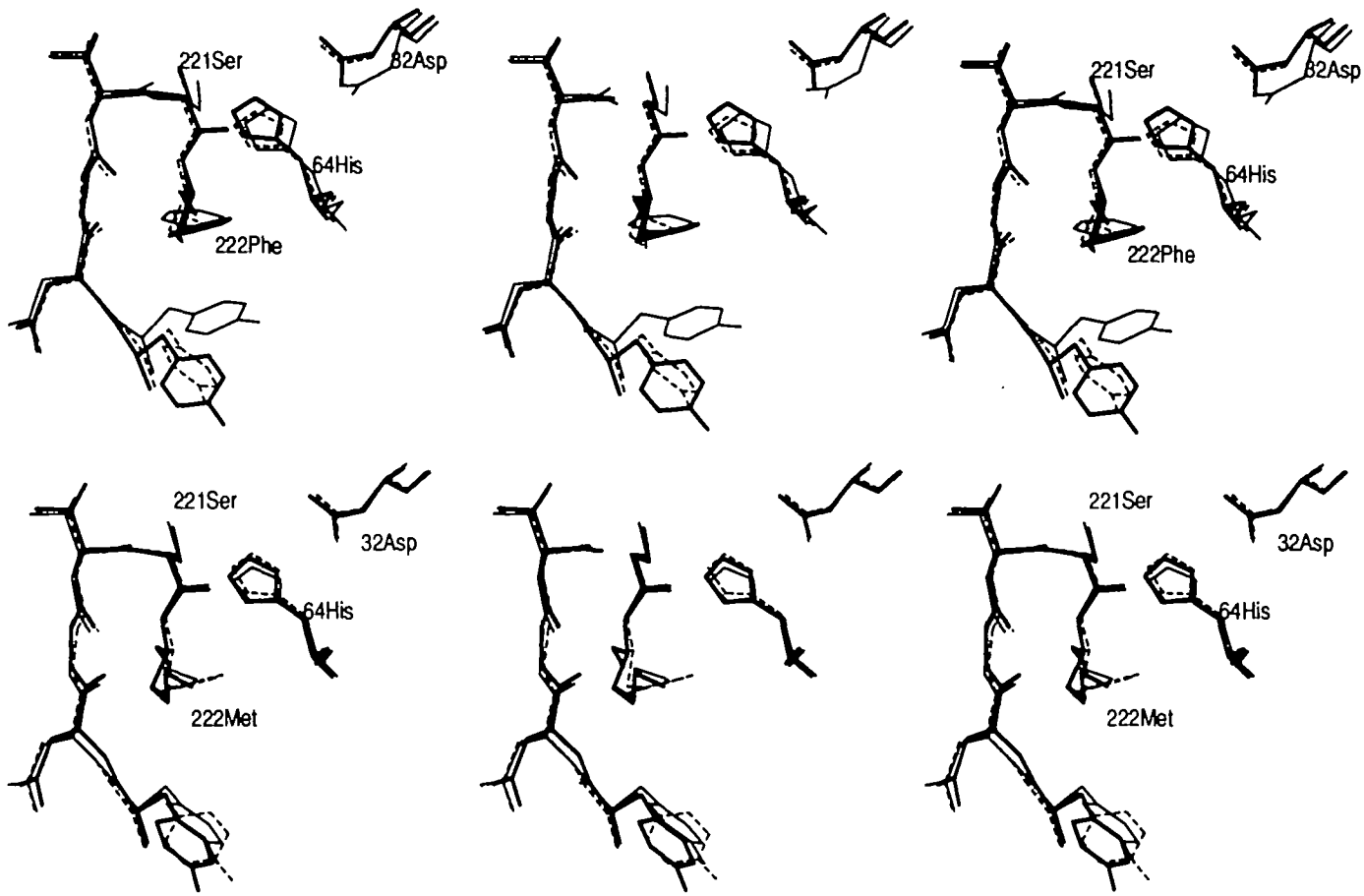

Fig. 7. (a) The average structure of the Phe mutant (bold lines, <Met(P1)Phe>). The average structure of the native enzyme (thin-dashed lines, <Met>) and the X-ray structure of the Phe-mutant (think lines) are shown for comparison. The residues His64, Ser221 and Phe222 are labelled. (b) The average structure of the Met mutant (bold lines, <Met(P1)Phe(P1)Met>) generated from the Phe-mutant is shown. The average structure Phe mutant (<Met(P1)Phe>) is shown in thin-dashed lines and the average native structure <Met> in thin lines. The residues His64, Ser221 and Met222 are labelled.

Tyr217 – phenol group rotates back to the original position, but at the end of the Phe → Met conversion the rotation is not completed yet; this explains the difference of 13° for Tyr217:χ_2 in the average structure <Met(P1)Phe(P1)met> relative to the average native structure.

The hydrogen bonding pattern in the mutant structure is somewhat different from the pattern in the native conformation. In the wild type enzyme, Met222:MN or Ala223:N forms a hydrogen bond to Gly219:O in a ratio 1:1. In the mutant enzyme the ratio of the two hydrogen bonds is 10:3. This may be induced by the increased bend between $\mu 3_{10}$-helix and α-helix, which makes the distance from Ala223:N to Gly219:O longer. In the wild type enzyme the oxy anion hole (Wells *et al.*, 1985) is oriented by a hydrogen bond from Thr220:N to Asn155:$O_{\delta 1}$, in the mutant enzyme the oxy anion hole is weakly oriented by a hydrogen bond from Thr220:$O_{\gamma 1}$ to a solvent molecule that has in turn a hydrogen bond to Asn155:$O_{\delta 1}$. When changing the phenylalanine into methionine again, the fixating hydrogen bond is not recovered. Though the original relative orientation of the peptide planes of Ser221 and Met222 is recovered ($\Delta\sigma \sim 0$, $\Delta\psi \sim 0$), the ratio of the hydrogen bond Met222:N → Gly219:O to Ala223:N → Gly219:O is 8:3, which indicates a Phe-like structure for the calculated Met-enzyme.

The predicted structural changes in this mutant are, as far as the main chain is concerned, in agreement with the X-ray structure. The rotation of the Tyr217 side chain is not observed in the X-ray structure. Considering the compactness of the S_1'-site in the mutant (B-factors of side chain of Tyr217 are 0.09 nm^2 in the mutant, 0.12 nm^2 in the wild type enzyme; B-factors of side chain of Phe222 are 0.03 nm^2) and the water channel located on the opposite side of the phenol group (Bott *et al.*, 1988; McPhalen and James, 1988), the change of the Tyr217 side chain conformation is plausible. The Phe-like hydrogen bonding patterns in the <Met(P1)Phe(P1)Met> structure (oxy anion hole fixation, hydrogen bonding ratio) is a neat demonstration of a situation in which the system relaxation time is considerably longer than the simulation period.

In a modelling study of this mutant (Bott, 1985), the shift of

Table VII. Quality of the predicted structure of the Met222Gln mutants

	X-ray	<Met(G1)Gln>
Dihedral (°)		
Ser221:χ_1	− 51	−11 (11)
Gln222:χ_1	− 79	−81 (8)
Gln222:χ_2	65	51 (13)
Gln222:χ_3	−130	−86 (14)
Distance (nm)		
His64:$N_{\epsilon 2}$ – Ser221:O_γ	0.327	0.430
Asn155:$N_{\delta 2}$ – Ser221:N	0.454	0.440
Ser221:O_γ	0.443	0.430
His64:$N_{\epsilon 2}$	0.700	0.672

Indicators of the quality of the predicted structure of the Met222Gln mutants. Averages over 10 ps.

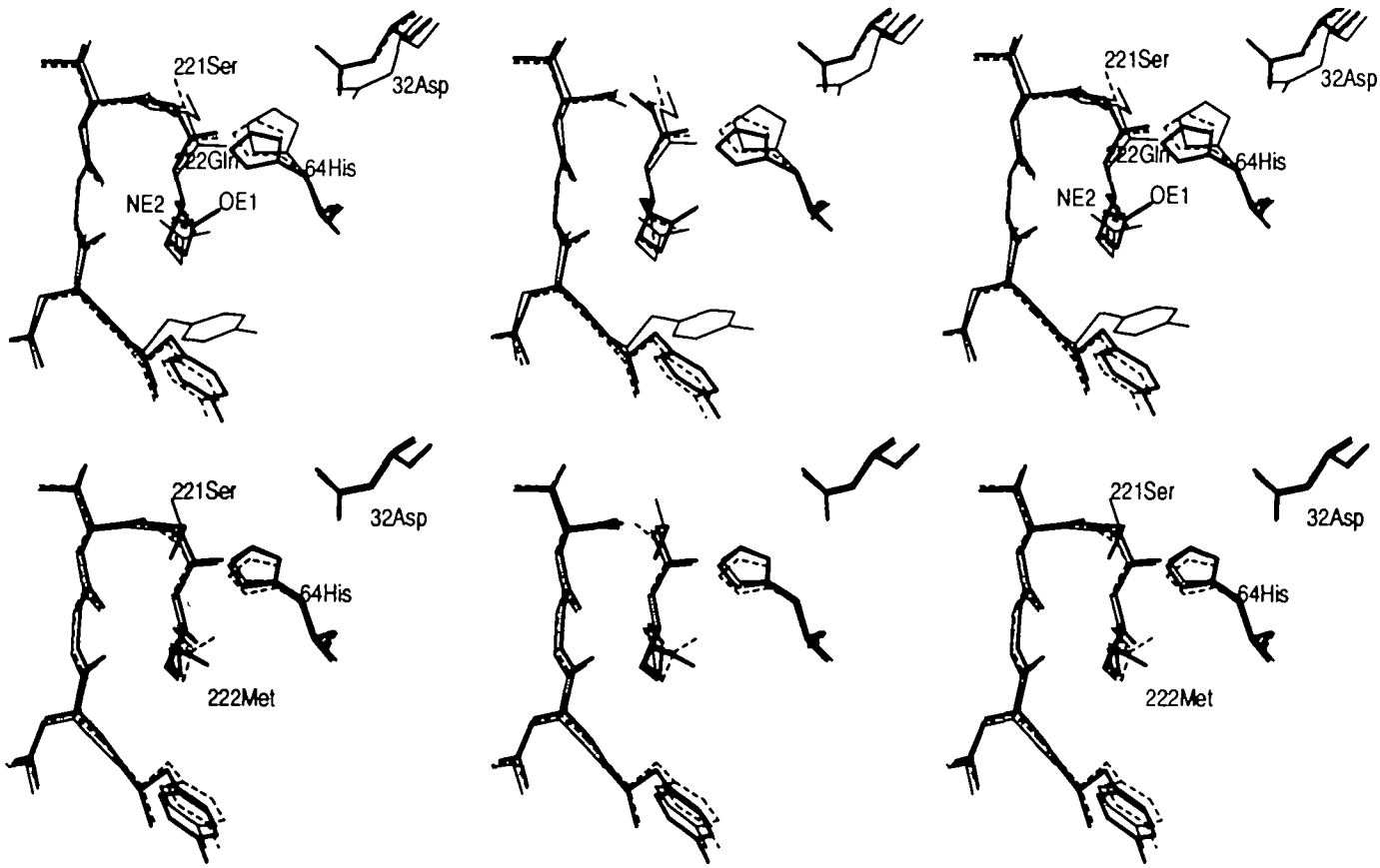

Fig. 8. (**a**). The average structure of the Gln-mutant (bold lines, <Met(G1)Gln>), the average structure of the native enzyme (thin-dashed lines, <Met>) and the X-ray structure of the Gln-mutant (thin lines) are shown. The residues His64 and Ser221 are labelled, as well as the oxygen of the amide group Gln222:$N_{\epsilon 2}$. (**b**) The average structure of the Met-mutant (bold lines, <Met(G1)Gln(G1)Met>) generated from the Gln-mutant is shown. The average structure of the Gln-mutant <Met(G1)Gln> is shown in thin-dashed lines and the average native structure <Met> is shown in thin lines. The residues His64, Ser221 and Met222 are labelled.

the backbone was not predicted and the modelled orientation of the phenyl plane was incorrect by 45°. Instead, a structural change of the active site was suggested because of the low specific activity of the mutant towards hydrolysis of sAAPF-pna. Recent studies revealed that the reduction of specific activity is due to the reduction of the space available to the P_1'-leaving group (Bonneau *et al.*, 1991). In the same study it was also found that for esters the catalytic rate is higher for the Phe-mutant enzyme than for the wild type enzyme. Since for this group of compounds deacylation is the rate-determining step, one may suggest that the observed angle between the 3_{10}-helix and α-helix reduces the ability to transfer protons (Hol *et al.*, 1978; Hol, 1985), thereby increasing the stability of the Michaelis complex over the tetrahedral intermediate in the deacylation step. Also, the distorted oxy anion hole may reduce the stability of the acylenzyme; both mechanisms may account for an increased overall catalytic rate.

However, the catalytic rate for esters of the mutant enzyme is only twice as fast as for the wild type enzyme and the structural explanations given are based on minor differences. As shown by Dijkman (1989), relatively small changes in the active site geometry can have a considerable effect on the calculated activation and stabilization energies. Structural changes of the activated enzyme relative to the free enzyme were established experimentally as well (James *et al.*, 1980). Hence, relating kinetic data to observed structural changes in the unbound enzyme is extremely hazardous; the best one can say is that the structural changes observed are not contradictory with the kinetic data available.

Mutant enzyme Met222 → Gln222

The average structure <Met(G1)Gln> obtained using the G1-protocol (Met222:C_ϵ → Gln222:$O_{\epsilon 1}$, create $N_{\epsilon 2}$, $H_{\epsilon 21}$, $H_{\epsilon 22}$ from dummy atoms) is very much the same as the average native structure (Figure 8a; r.m.s.PD of the backbone atoms 0.025 nm, r.m.s.PD of the side chain atoms 0.038 nm). The largest difference is found for the active site serine; the χ_1-angle has rotated over −28° (cf. Tables IV and VII). The side chain dihedral angles of the glutamine agree well with the values derived from the X-ray structure, except for the χ_3-angle; a difference of −144° is observed (Table VII, Figure 8a). The path chosen may force the terminal amide group into a conformation with higher energy than one of its rotamers. To see whether a different path would give better agreement with the experimental structure, we used the alternative atomic map G2 (Met222:$C_{\epsilon 2}$ → Gln222:$N_{\epsilon 2}$, create $O\epsilon_1$, $H_{\epsilon 2}$, $H_{\epsilon 22}$ from dummy atoms). After 2 ps simulation the temperature had increased by a factor 3 and after 8 ps simulation the leap-frog algorithm did not converge any more. This indicates that using the G2 atomic map the Gln222 side chain amide dipole is oriented against the local electric field; steric hindrance prevents the amide dipole orienting parallel to the electric field.

Performing the backward conversion, the conformation of the methionine side chain in the <Met(G1)Gln(G1)Met> structure agrees well with the conformation in the native structure (Figure 8b, Table IV). The r.m.s.PD of both the backbone atoms and side chain atoms with respect to the average native structure is small (0.018 nm and 0.036 nm respectively). The major

difference is the Ser221:χ_1 conformation that is gauche$^-$; the dominant wild type conformation is gauche$^+$ (Table IV).

Since glutamine is a polar residue, the hydrogen bonding pattern near the active site is expected to be somewhat different from the wild type enzyme. The side chain amide nitrogen is fixed by hydrogen bonds to Asn218:O and His64:$N_{\epsilon 2}$ and replaces water molecule S2 (Figure 5a) in the wild type enzyme. This water molecule hops to the bottom of the S_1' binding site where it becomes hydrogen bonded to Gln22:$O_{\epsilon 1}$ and His67:$N_{\epsilon 1}$. The structure and fixation of the oxy anion hole remains intact. The nucleophilic hydroxyl group loses the hydrogen bond Ser125:O and forms one with Asn155:$N_{\epsilon 2}$.

In the average structure <Met(Gl)Gln(G1)Met> the wild type hydrogen bonding pattern is not restored; in addition, a hydrogen bond from Ser221:O_γ to the imidazole group is observed.

Whether the predicted conformation of the glutamine side chain is correct is hard to judge. The orientation of the amide plane is off by 36°; the assignment of the atoms in the X-ray structure makes the difference increase to 144°, but the calculated orientation of the angle between the electric field due to the protein atoms and amide dipole is rather favourable (8°) in the MD structure.

The assignment of the respective positions in the X-ray structure is based on a minor (structural) shift of the imidazole group away from the glutamine side chain (Bott and Ultsch, 1986) and the reduction of the specific activity of the enzyme for hydrolysis of sAAPF-pna. In this model it is assumed that the electrostatic field of Gln222:$O_{\epsilon 1}$ destabilizes the transition state. This destabilization of the transition state will also occur when the dipolar field is reversed. Another explanation to account for the reduced activity of the Gln−enzyme is the formation of the hydrogen bond from Gln222:$N_{\epsilon 2}$, which reduces the rotational freedom of the imidazole. It has been argued that during the acylation step (rate-determining for hydrolysis of sAAPF-pna) the imidazole rotates to form the tetrahedral intermediate (Drenth *et al.*, 1975; van Duynen *et al.*, 1979). The motional freedom of the latter needed in the formation of the tetrahedral intermediate is restricted by the Gln222 side chain dipole field. A definite assignment may be obtained from, e.g. neutron diffraction data.

The insertion of the polar side chain causes a massive solvent rearrangement; the water molecule bound to the oxy anion hole moves over several Å, even during the equilibration period. This is probably due to the chosen functional form of $\lambda(t)$. During the last ps of the conversion the driving force for this process (the Coulomb interaction) increases from 0.53 to 1.0. Another example of slow electrostatic equilibration is observed in the <Met(G1)Gln(G1)Met> structure; the hydrogen bond from Ser221:O_γ to His64:$N_{\epsilon 2}$ is typical for the Michaelis complex, not for the free enzyme. The origin of this hydrogen bond is probably that the Gln222:$N_{\epsilon 2}$ drags the imidazole towards the nucleophilic hydroxyl. When the glutamine side chain is converted into methionine the hydrogen bond is transferred to Ser221:O_γ. The absence of the hydrogen bond from Ser221.O_γ to Ser125:O also indicates that the <Met(G1)Gln(G1)Met> structure is a strained structure.

Conclusions

In this study a modelling technique derived from the free energy perturbation methodology is used to predict the structure of mutants of subtilisin BPN' at position Met222. The method is based on the slow growth procedure. The emphasis is on 'near equilibrium conversion' rather than on sufficient sampling of the conformational space along the path of conversion. This allows for a relatively fast change of one mutant into the other. Unfortunately the test for a 'near equilibrium conversion' is now much harder, since a check on a zero hysteresis of the free energy difference is made impossible. Instead, we used a combination of comparison of calculated structure with the X-ray structure, hydrogen bonding patterns and changes in activity to assess the validity of the structural prediction.

The structural changes can be divided into two groups; structural changes of atoms for which interaction parameters change (first order changes) and structural changes induced in the environment (second order changes). The latter may range from tiny rotations due to van der Waals repulsion to diffusion of solvent molecules into sites left empty by first order changes.

The first class of changes is relatively easy to predict; a simple substitution of the residue may already reveal the preferred conformation; the conformation is mainly determined by packing criteria. However, details of the conformation, as compared to the X-ray structure, may be better predicted by the slow growth procedure. When changing the mutant residue back into the wild type methionine, the conformation of the latter was correct in all cases, no matter how tense the conformation of nearby residues was.

Induced changes are much harder to predict using replacement techniques only. As was illustrated in the conversion of methionine into phenylalanine, the conformational change of the tyrosine side chain was not anticipated in the modelling study, though it is a realistic option. Also the gauche$^-$ conformation in the Ala-mutant was rather unexpected. The latter conversion also illustrates that a certain protocol may yield the correct mutant structure, but that a successful prediction depends very much on the initial configuration. Only when the conversion is made very slowly, do we expect that this configurational dependency of the predicted structure will vanish.

Whether or not the conversion was made 'near equilibrium' was tested by comparing the average hydrogen bonding pattern in the native simulation and the Met()Y()Met simulation. This appeared to be a very sensitive tool: in almost all cases the pattern depended on the history of the simulation, whereas the r.m.s.PD did not reveal this. This may be caused by the 'polarization feedback' of the Coulomb interaction. Also, since the functional form of $\lambda(t)$ was made optimal for the van der Waals interaction, the Coulomb interaction changed 50% during the last 2 ps of the conversion. A linear time dependency for this interaction would be more optimal. Another strategy to reduce the dependency on the history is to reduce the overall interaction along the path of slow growth (Mark *et al.*, 1991). In order to obtain reliable free energy differences (i.e. to sample correctly), the conversion should be made orders of magnitude slower.

Finally, the structural changes were compared with the kinetic data available. In all cases the changes were compatible with the data, although we had to make a few assumptions to explain the activity of the Ala-mutant relative to the wild type enzyme. However, the kinetic activity is very much dependent on minor structural details, which makes it hazardous to rely on structure−activity relations too heavily.

Acknowledgements

The authors gratefully thank Dr R.R.Bott of Genencor Inc. for providing us with the coordinates of native subtilisin BPN' as well as the mutants prior to publication and thank him for helpful discussions and a reading of the manuscript. The financial support of Genencor Inc. is gratefully acknowledged.

References

Alber,T., Doa-Pin,S., Wilson,K., Wozniak,J.A., Cook,S.P. and Matthews,B W. (1980) *Nature*, **330**, 41–41.

Berendsen,H.J.C., Postma,J.P.M., van Gunsteren,W F., DiNola,A and Haak,J.R (1984) *J. Chem. Phys.*, **81**, 3684–3690.

Berendsen,H.J.C., Postma,J.P.M. and van Gunsteren,W.F. (1985) In Hermans,J. (ed.), *Molecular Dynamics and Protein Structure*. PolyCrystal Book Service, Western Springs, IL, USA, pp. 43–46.

Berendsen,H.J.C., van Gunsteren,W.F., Zwinderman,H.R.J. and Geurtsen,R.G (1986) *Ann. NY Acad. Sci.*, **482**, 269–285.

Berendsen,H.J.C. (1991) In Renugopalakrishnan,V., Cary,P.R., Smith,I C.P., Huang,S.G and Storer,A.C. (eds), *Proteins: Structure, Dynamics, and Design*. Escom, Leiden, The Netherlands, pp 384–392.

Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F., Brice,M.D., Rodger,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542.

Bonneau,P.R , Graycar,T.P., Estell,D.A. and Jones,J.B. (1991) *J. Am. Chem. Soc.*, **113**, 1026–1030.

Bott,R.R. (1985) *World Biotech. Rep..*, **2**, 51–59.

Bott,R.R., Ultsch,M., Wells,J., Powers,D., Brudick,D , Struble,M., Burnier,J., Estell,D., Miller,J., Graycar,T., Adams,R. and Power,S. (1985) In LeBaron,H.M., Mumma,R.O., Honeycott,R.C. and Duesing,J.H. (eds), *Biotechnology in Agricultural Chemistry*. ACS Symposium Series 334, Washington DC, USA, pp. 139–147.

Bott,R.R. and Ultsch,M. (1986) In Alacevic,M., Hranueli,D. and Toman,Z. (eds), *Proc. Fifth Int. Symp. Genetics of Industrial Microorganisms*, Zagreb, pp. 375–385.

Bott,R.R., Ultsch,M., Kossiakoff,A., Graycar,T., Katz,B. and Power,S (1988) *J. Biol. Chem.*, **263**, 7895–7906.

Craik,C.S., Largman,C., Fletcher,T., Roczniak,S., Barr,P.J., Fletterick,R. and Rutter,W.J. (1985) *Science*, **228**, 291–297.

Dijkman,J.P. (1989) PhD Thesis, University of Groningen, The Netherlands.

Doolittle,R.F. (1981) *Science*, **214**, 149–159.

Drenth,J., Swen,H.M., Hoogenstraaten,W and Sluyterman,L.E.A.E. (1975) *Proc. Koninkl. Akad. Wetensch.*, **C78**, 104–110.

Estell,D.A., Graycar,T.P. and Wells,J.A. (1985) *J. Biol. Chem.*, **262**, 6518–6521.

Estell,D.A., Graycar,T.P., Miller,J.V., Powers,D.B., Burnier,J.P., Ng,P.G. and Wells,J.A. (1986) *Science*, **233**, 659–662.

Geller,M., Swanson,S.M. and Meyer,E.F.,Jr (1990) *J. Biomol. Struct. Dynamics*, **7**, 1043–1052.

Geller,M., Swanson,S.M. and Meyer,E.F.,Jr (1991) *J. Am. Chem. Soc.*, **112**, 8925–8931.

Graycar,T.P. (1991) In Dordick,J.S. (ed.) *Biocatalysis in Industry*. Plenum, New York, NY, pp. 237–283.

Heiner,A.P., Berendsen,H.J.C. and van Gunsteren,W.F. (1992) *Proteins: Struct. Funct. Genet.*, **14**, 451–464.

Hermans,J., Yun,R.H. and Anderson,A.G. (1992) *J. Comput. Chem.*, **13**, 429–442.

Hol,W.G.J., Van Duynen,P.T. and Berendsen,H.J.C. (1978) *Nature*, **273**, 443–446.

Hol,W.G.J. (1985) *Prog. Biophys. Mol. Biol.*, **45**, 149–195.

James,M.N.G., Sielecki,A.R., Brayer,G.D., Delbaere,L.T.J. and Bauer,C.-A. (1990) *J. Mol. Biol.*, **144**, 43–88.

Janin,J., Wodak,S., Levitt,M. and Maigret,B. (1978) *J. Mol. Biol.*, **125**, 357–386.

Karplus,M. and Petsko,G.A. (1990) *Nature*, **347**, 631–637.

Köhler,J., Sänger,W. and van Gunsteren,W.F. (1987) *Eur. Biophys. J.*, **15**, 197–210.

Lee,C. and Subbiah,S. (1991) *J. Mol. Biol.*, **217**, 373–388.

Mark,A.E., van Gunsteren,W.F. and Berendsen,H.J.C. (1991) *J. Chem. Phys.*, **94**, 3808–3816.

Mazor,M. and Pettitt,B.M. (1991) *Mol. Sim.*, **6**, 1–4.

McPhalen,C.A. and James,M.H.G. (1988) *Biochemistry*, **27**, 6582–6598.

Meyer,E.F., Cole,G.R., Radhakrishnan,R. and Epp,O. (1988) *Acta Crystallogr.*, **B44**, 26–38.

Ponder,J.W. and Richards,F.M. (1987) *J. Mol. Biol.*, **193**, 775–791.

Postma,J.P.M. (1985) PhD Thesis, University of Groningen, The Netherlands.

Ryckaert,J.-P., Ciccotti,G. and Berendsen,H.J.C. (1977) *J. Comput. Phys.*, **23**, 327–341.

Schechter,I. and Berger,A. (1967) *Biochem. Biophys. Res. Commun* , **27**, 157–162.

Sielecki,A R , Hendrickson,W.A , Boroughton,C.G., Delbaere,L.T.J., Brayer,G.D. and James,M.N.G. (1979) *J Mol. Biol.*, **134**, 781–804

Stauffer,C.E. and Etson,D. (1969) *J. Biol. Chem.*, **244**, 5333–5338.

Teleman,O. and Jönsson,B. (1986) *J. Comput. Chem.*, **7**, 58–66.

Thomas,P.G., Russle,A.J. and Fersht,A.R. (1985) *Nature*, **318**, 375–376.

van Duynen,P.Th., Thole,B.T. and Hol,W.G. (1979) *J. Biophys. Chem.*, **9**, 273–280.

van Gunsteren,W.F. and Berendsen,H.J.C. (1977) *Mol. Phys.*, **34**, 1311–1327.

van Gunsteren,W.F. and Berendsen,H.J.C. (1987) In van Gunsteren,W.F. and Berendsen,H.J.C. (eds), *Groningen Molecular Simulation (GROMOS) Library Manual*. Biomos, Nijenborgh 16, Groningen, The Netherlands.

van Gunsteren,W F. (1988) *Protein Engng*, **2**, 5–13.

van Gunsteren,W.F. (1989) In van Gunsteren,W.F. and Weiner,P (eds), *Computer Simulation of Biological Systems*. Escom, Leiden, pp. 27–59.

van Gunsteren,W.F. and Berendsen,H.J.C. (1990a) *Angew. Chem. Int. Ed. Engl.*, **29**, 992–1023.

van Gunsteren,W.F. (1990b) In Troyanowsky,C. (ed.), *Modelling of Macromolecular Structures and Properties*. Elsevier, Amsterdam.

Wells,J.A., Cunningham,B.C., Graycar,T P. and Estell,D.A. (1985) *Phil. Trans. R. Soc. London A*, **317**, 415–423.

Wells,J.A. (1990) *Biochemistry*, **29**, 8509–8517.

Received on June 23, 1992; revised on October 27, 1992; accepted on January 30, 1993